SUPPLEMENTARY INFORMATION

Hs-HBD purification

Wild-type and mutant Hs-HBD were purified on a 10 ml Ni column. *E. coli* cells were by sonication in buffer A (40 mM NaH_2PO_4 (pH 7.0), 1 M NaCl, 5% glycerol, 2.8 mM β-mercaptoethanol and 10 mM imidazole) with protease inhibitors (Sigma-Aldrich). After the cleared cell lysate was applied, the Ni column was washed with 10 column volume (CV) (100ml) buffer A plus 60 mM imidazole. The protein was eluted with a linear gradient of 60 to 300 mM imidazole in 15 CV. The protein peak was collected and exchanged by dialysis to B buffer (40 mM NaH_2PO_4 (pH 7.0), 100 mM NaCl, 5% glycerol, 2 mM DTT, and 0.5 mM EDTA) for thrombin digestion. The proteins were further purified on a Mono S column (GE Healthcare) with a 100 mM - 500 mM gradient of NaCl in buffer B. Fractions containing Hs-HBD were concentrated in the storage buffer (20 mM HEPES (pH 7.0), 75 mM NaCl, 5% glycerol, 0.5 mM EDTA, 2 mM DTT) and stored at 4 or -20°C. Selenomethionine-labeled (SeMet) Hs-HBD V71M was expressed in methionine-auxotrophic *E. coli* strain B834 (DE3) in a defined medium (Hendrickson et al., 1990) and purified using the same protocol.

Structure determination and analysis

The crystals of Hs-HBD complexed with the 12bp RNA/DNA belonged to $P2_12_12_1$ space group. They diffracted X-rays to 2.4 Å resolution but with high mosaicity and irregular spot shape. Hs-HBD contains only one natural Met (M27) at the N-terminus, which is potentially disordered. Thus, one (position 9), two (4 and 9), or three (4, 9 and 11) 5-iodouridines were introduced into the DNA strand of the RNA/DNA hybrid for phasing by isomorphous replacement (see supplementary Table I). Crystals were grown

with these nucleic acids and diffraction data were collected. The single substitution (position 9) resulted in crystals that diffracted X-rays to 2.1 Å resolution, but the structure could not be solved by SIR. This may be due to two reasons: low occupancy of iodine (refined to 0.5 after the structure was solved) possibly due to loss during oligo purification or diffraction data collection, and non-isomorphism between the two crystals as indicated by the diffraction qualities.

To solve the structure by SeMet replacement, each of the four Val residues was systematically replaced by Met. Only the V71M substituted protein could be expressed in a soluble form. Crystals of SeMet-substituted V71M mutant HBD complexed with the 12bp RNA/DNA were grown, but they did not diffract X-rays. RNA/DNA hybrids of various lengths were tried next in combination with the SeMet V71M protein, and co-crystals of a 6bp RNA/DNA were obtained. A SAD dataset was collected at Advanced Photon Source ID-22 beamline at 0.97928 Å (Se absorption peak) to 2.4 Å resolution and processed in $P3_221$ space group in HKL2000 (Otwinowski, 1997). SOLVE (Terwilliger and Berendzen, 1999) was used to locate 2 selenium sites from the single protein (N-terminal M27 and M71) molecule in an asymmetric unit. The experimental maps were traced in COOT (Emsley and Cowtan, 2004) and O (Jones et al., 1991) to produce a model of the complex containing one protein molecule and one 6bp RNA/DNA. The asymmetric unit also contained a half of RNA/DNA with crystallographic two-fold axis going through it. To avoid the problem with the two-fold axis, two copies of the initial model were then refined in a lower symmetry space group - $P3_2$ using CNS (Brunger et al., 1998) interspersed with manual building. The asymmetric unit contained two protein molecules and three RNA/DNA hybrids. Two were binding to

the genuine RNA/DNA interfaces and one was forming the crystal lattice by bridging two protein molecules by interactions between RNA/DNA ends and protein surface opposite to the genuine RNA/DNA interface. Refinement of this model ended with exceedingly high R values most likely due to a disorder in the third RNA/DNA molecule, which had poor electron density and high B-factors. The partially refined model, however, was used successfully in molecular replacement to solve the structure of the 12bp RNA/DNA complex containing a single 5-iodouridine substitution. The complete 12-mer complex structure containing two RNA/DNAs and six protein molecules was built in COOT and O and refined with CNS (Table I). The protein molecules could be traced from the N-terminus to residues 72 to 74, depending on the protein molecules.

The nucleic acid conformation was analyzed using CURVES (Lavery and Sklenar, 1988). Structure analyses including superpositions were done in O or MolMol (hugin.ethz.ch/wuthrich/software/molmol). Surface potentials were calculated with GRASP (Nicholls et al., 1991). Figures were prepared using PyMol (www.pymol.org).

Sedimentation equilibrium

Solution densities ρ were measured at 20°C on a Mettler-Toledo DE51 density meter and corrected to values for ρ at 4°C. The partial specific volume v for the Hs-HBD was calculated based on the amino acid composition using SEDNTERP 1.09 (www.jphilo.mailway.com), as was the extinction coefficient. Extinction coefficients at 280 nm for the double stranded nucleic acids were based on the extinction coefficients of the single strands at 260 nm and spectrophotometric measurements of the double stranded species.

Characterization of the 12bp nucleic acids. Sedimentation equilibrium experiments on the double stranded 12bp nucleic acids were carried out at 16, 19, 25, 28, 31 and 33 krpm and with loading volume of 130 μL at concentrations of 0.7 – 0.9 A_{280}. A global analysis in terms of a single ideal solute showed excellent fits and in all cases data were consistent with the presence of a monomeric double stranded species (See Table below). The partial specific volumes calculated based on the experimental buoyant molecular masses range from 0.53 – 0.56 cm^3g^{-1}, and are within the range expected for nucleic acids. The experimentally determined buoyant molecular masses were used in subsequent calculations.

12-mer nucleic acid	Experimental M(1 – νρ) / Da	Calculated M / Da	M(1 – νρ) / Da ν = 0.55 cm^3g^{-1}	Stoichiometry ν = 0.55 cm^3g^{-1}	Experimental ν / cm^3g^{-1}
RNA/DNA hybrid	3400 ± 30	7440.8	3272	1.04 ± 0.01	0.533 ± 0.004
dsRNA	3300 ± 50	7590.7	3338	0.99 ± 0.01	0.555 ± 0.007
dsDNA	3300 ± 30	7290.9	3206	1.03 ± 0.01	0.537 ± 0.004

The RNA/DNA hybrid forms a high affinity 1:1 complex with HBD. Sedimentation equilibrium experiments were carried out at 3:1 and 1:1 protein : nucleic acid molar ratios and three loading concentrations of 1.7, 3.4 and 5.1 μM RNA/DNA hybrid. Sedimentation equilibrium was achieved within 48 hours. The data collected for the 3:1 ratio were not further analyzed due to sample losses observed with increasing rotor speeds. In light of the crystal structures, the aggregation could be due to crosslinking of nucleic acids by HBD as observed in the 12bp structure (Figure 2A), crosslinking of HBDs by the hybrid duplex as observed in the 6bp complexes, or both.

The 1:1 protein:nucleic acid samples were, however, well behaved and yielded excellent fits when analyzed in terms of a single ideal solute (Supplementary Figure 1), returning a buoyant molecular mass of 4970 ± 100 Da. Using the amino acid sequence of the HBD and the experimental buoyant molecular mass for the nucleic acid duplex, the 1:1 complex is expected to have a buoyant molecular mass of 5057 Da. The experimental data therefore support complete 1:1 complex formation with a stoichiometry of 0.98 ± 0.02. Similar experiments were carried out at three lower loading concentrations of the 1:1 complex, corresponding to 1.0, 2.0 and 3.0 μM, with data collected at 230 nm. Data were again consistent with the presence of a single ideal solute having a mass corresponding to that expected for the 1:1 complex (data not shown). As 1:1 complex formation is complete at all loading concentrations studied, the interaction between HBD and the hybrid duplex is characterized by a $K_{d,1} \leq 0.2$ μM.

The dsRNA interacts with HBD to form 1:1 and 2:1 complexes. Sedimentation equilibrium experiments carried out at a 3:1 protein : nucleic acid molar ratio and loading concentrations of 1.6, 3.1 and 4.9 μM were consistent with complex formation, returning a weight average buoyant molecular masses at least equivalent to that expected for the 1:1 complex. Attempts to model the data in terms of a reversible A + B ⇆ AB equilibrium using mass conservation resulted in poor data fits. However, excellent fits were obtained when the data were analyzed in terms of a reversible A + B + B ⇆ AB + B ⇆ ABB equilibrium in which both 1:1 and 2:1 HBD : dsRNA complexes are formed (Supplementary Figure 2). Sedimentation equilibrium experiments carried out at a 1:1 protein:nucleic acid ratio (loading concentrations of 1.5, 3.0 and 4.6 μM) were also consistent with reversible complex formation and a global analysis of both the 3:1 and

1:1 data sets in terms of a reversible A + B + B ⇆ AB + B ⇆ ABB equilibrium yielded values of $K_{d,1}$ = 4.9 ± 0.5 μM and $K_{d,2}$ of 6.8 (-1.9, + 0.6) μM with excellent fits (Supplementary Figure 2).

The dsDNA interacts with HBD to form 1:1 complexes. Sedimentation equilibrium experiments carried out at a 1:1 protein : nucleic acid molar ratio and concentrations of 1.5, 3.0 and 4.8 μM returned weight average buoyant molecular masses just larger than those observed for the free double stranded 12-mer. When data were modeled in terms of a reversible A + B ⇆ AB equilibrium, using mass conservation, excellent fits were observed (Supplementary Figure 3). This analysis returns a $K_{d,1}$ of 25 ± 5 μM, consistent with the overall weaker binding observed in a qualitative data analysis (Figure 1A). Experiments carried out using a 3:1 molar ratio were analyzed in terms of the same model. Mass conservation and constraining the protein : DNA ratio to 3 did not result in satisfactory data fits. Excellent fits, however, were observed in the absence of these constraints, resulting in a $K_{d,1}$ of 23 ± 5 μM (data not shown). Additional experiments will be needed to determine the distributions of 1:1 and 2:1 complexes and the $K_{d,2}$ of AB + B ⇆ ABB equilibrium,

REFERENCES

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**: 905-921.

Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132.

Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**: 1665-1672.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**: 110-119.

Lavery, R. and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J Biomol Struct Dyn* **6**: 63-91.

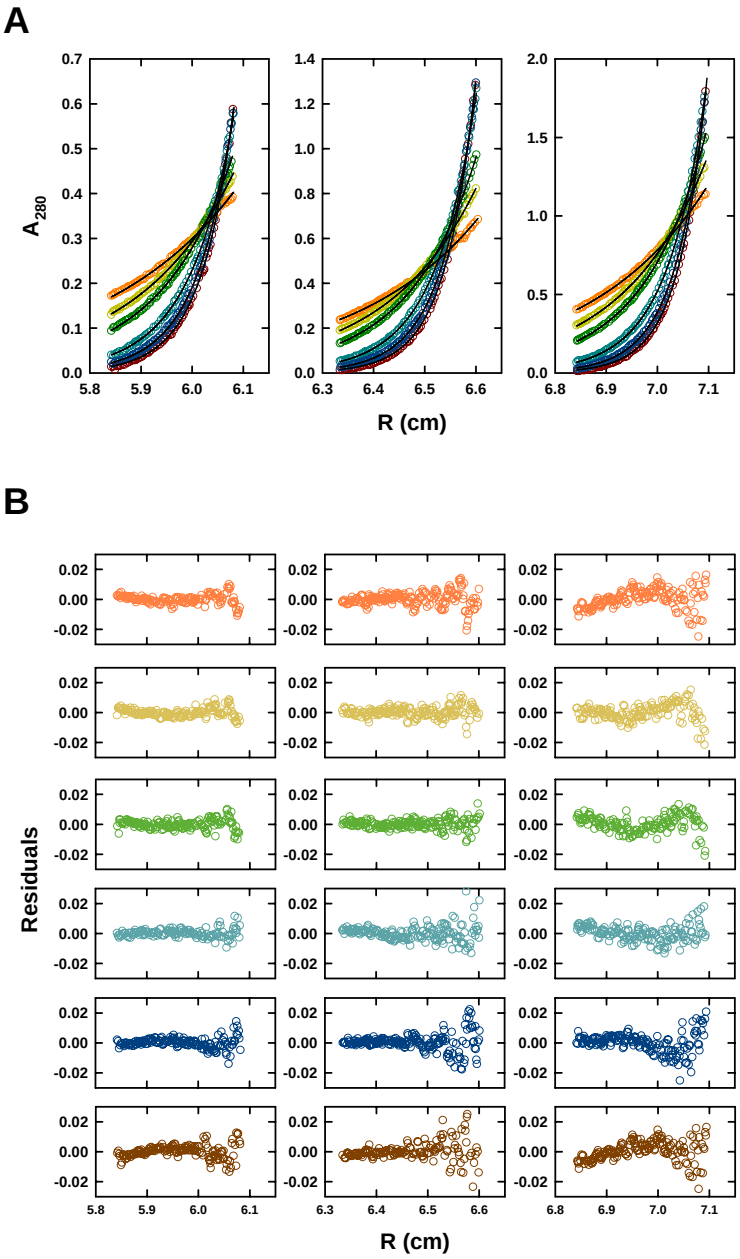
Nicholls, A., Sharp, K.A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**: 281-296.

Otwinowski, Z., Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymol*, Carter, C.W., Sweet, R. M. (ed.), Vol. 276, pp. 307-326. New York: Academic Press

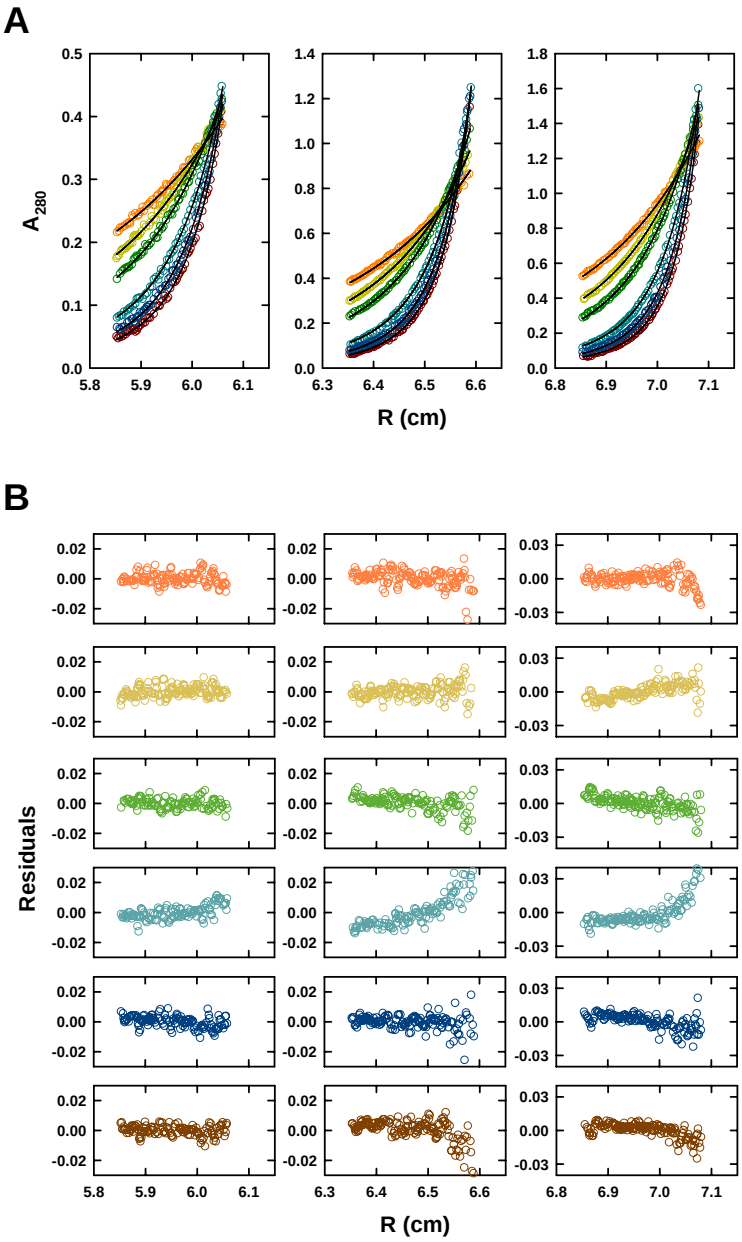
Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55 (Pt 4)**: 849-861.

Supplementary Table I. Sequences of the oligonucleotides

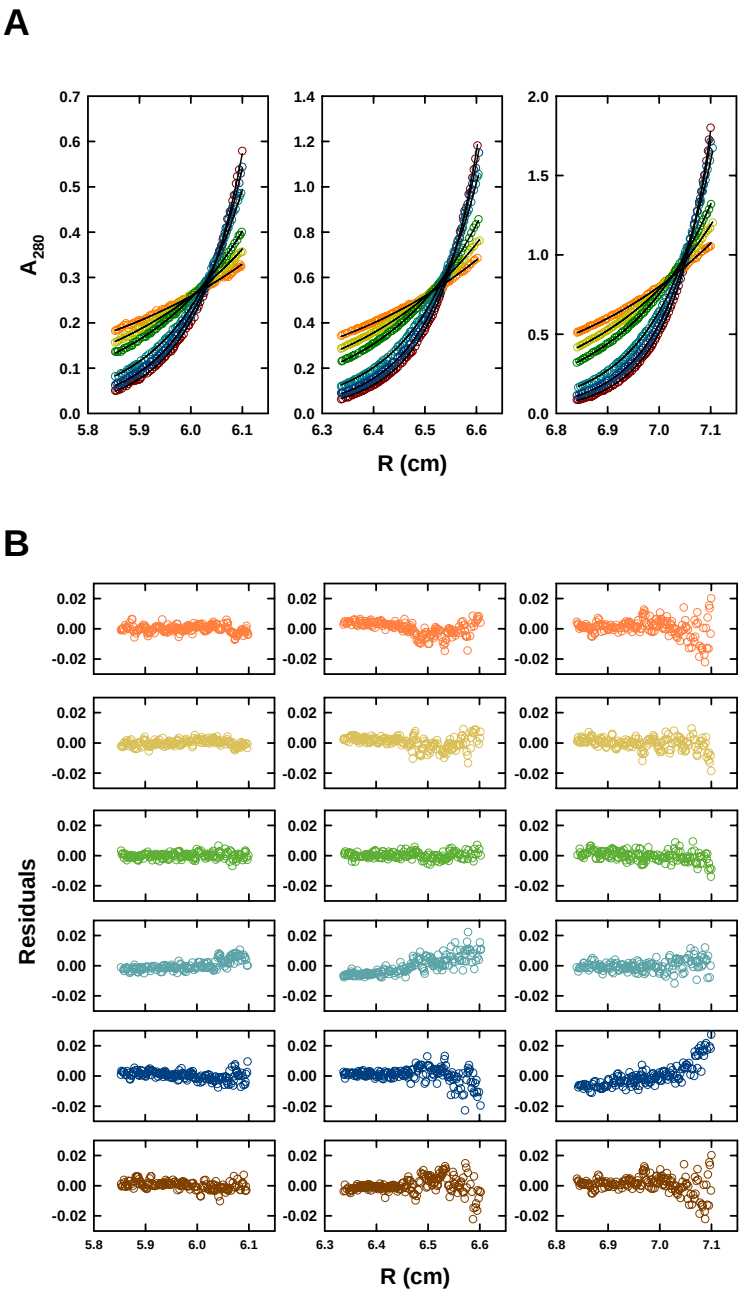
	Strand 1	**Strand 2**
6bp RNA/DNA	r(ACAUCG)	d(CGATGT)
12bp RNA/DNA	r(GACACCUGAUUC)	d(GAATCAGGTGTC)
12bp RNA/DNA (single iodouridine, position 9)	r(GACACCUGAUUC)	d(GAATCAGG(5-IU)GTC)
12bp RNA/DNA (two iodouridines, positions 4 and 11)	r(GACACCUGAUUC)	d(GAA(5-IU)CAGGTG(5-IU)C)
12bp RNA/DNA (three iodouridines, positions 4, 9 and 11)	r(GACACCUGAUUC)	d(GAA(5-IU)CAGG(5-IU)G(5-IU)C)
12bp RNA/RNA	r(GACACCUGAUUC)	r(GAAUCAGGUGUC)
12bp DNA/DNA	d(GACACCTGATTC)	d(GAATCAGGTGTC)
36bp RNA/DNA	r(UGGGGGCUCGUCCGGGAUAUGGG AACCACUGAUCCC)	d(GGGATCAGTGGTTCCCATATC CCGGACGAGCCCCCA)



Supplementary Figure 1. HBD and the RNA/DNA hybrid form a high affinity 1:1 complex. (**A**) Sedimentation equilibrium profiles obtained for the 1:1 HBD to 12-bp RNA:DNA hybrid shown in terms of A_{280} versus the radius r for data collected at loading A_{280} of 0.26 (**left**), 0.51 (**center**) and 0.80 (**right**). Data were collected at 4.0°C and rotor speeds of 16 (orange), 19 (yellow), 22 (green), 28 (cyan), 31 (blue) and 33 (brown) krpm and analyzed globally in terms of a single ideal solute. Best fits, corresponding to a monodisperse 1:1 complex, are shown as black lines through the experimental points. For clarity, alternate data points have been omitted. (**B**) The corresponding distributions of the residuals.

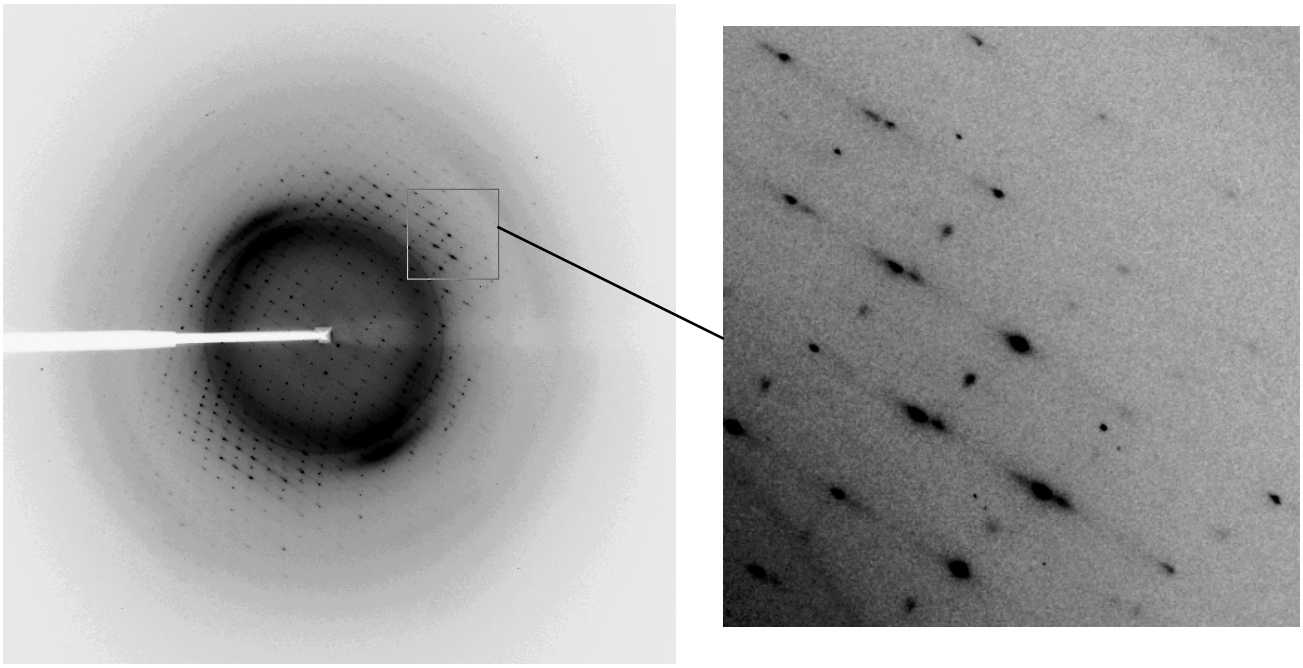


Supplementary Figure 2. HBD and dsRNA form 1:1 and 2:1 complexes. (**A**) Sedimentation equilibrium profiles obtained for the 3:1 HBD to 12-bp dsRNA shown in terms of A_{280} versus the radius r for data collected at a loading A_{280} of 0.33 (**left**), 0.65 (**center**) and 1.04 (**right**). Data were collected at 4.0°C and rotor speeds of 16 (orange), 19 (yellow), 22 (green), 28 (cyan), 31 (blue) and 33 (brown) krpm and analyzed globally (along with the 1:1 data not shown) in terms of a reversible A + B + B ⇆ AB + B ⇆ ABB equilibrium (A = dsRNA). Best fits are shown as black lines through the experimental points. For clarity, alternate data points have been omitted. (**B**) The corresponding distributions of the residuals.

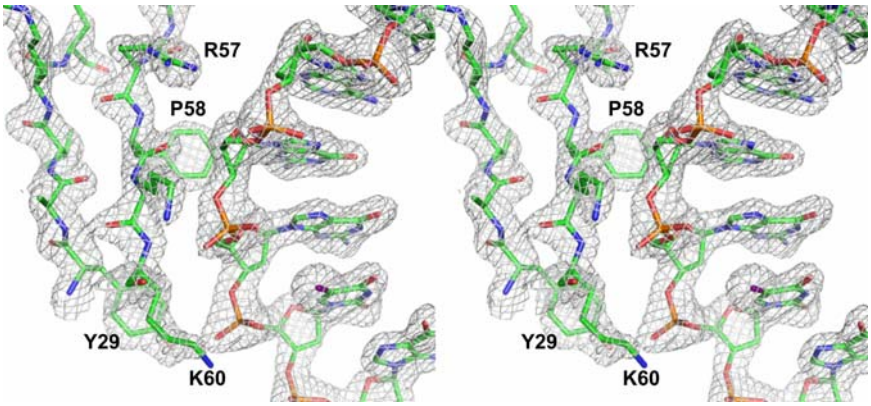


Supplementary Figure 3: HBD and dsDNA form a low affinity complex. (**A**) Sedimentation equilibrium profiles obtained for the 1:1 HBD to 12-bp dsDNA shown in terms of A_{280} versus the radius r for data collected at a loading A_{280} of 0.22 (**left**), 0.46 (**center**) and 0.72 (**right**). Data were collected at 4.0°C and rotor speeds of 16 (orange), 19 (yellow), 22 (green), 28 (cyan), 31 (blue) and 33 (brown) krpm and analyzed globally in terms of a reversible A + B ⇆ AB equilibrium (A = dsDNA). Best fits are shown as black lines through the experimental points. For clarity, alternate data points have been omitted. (**B**) The corresponding distributions of the residuals.

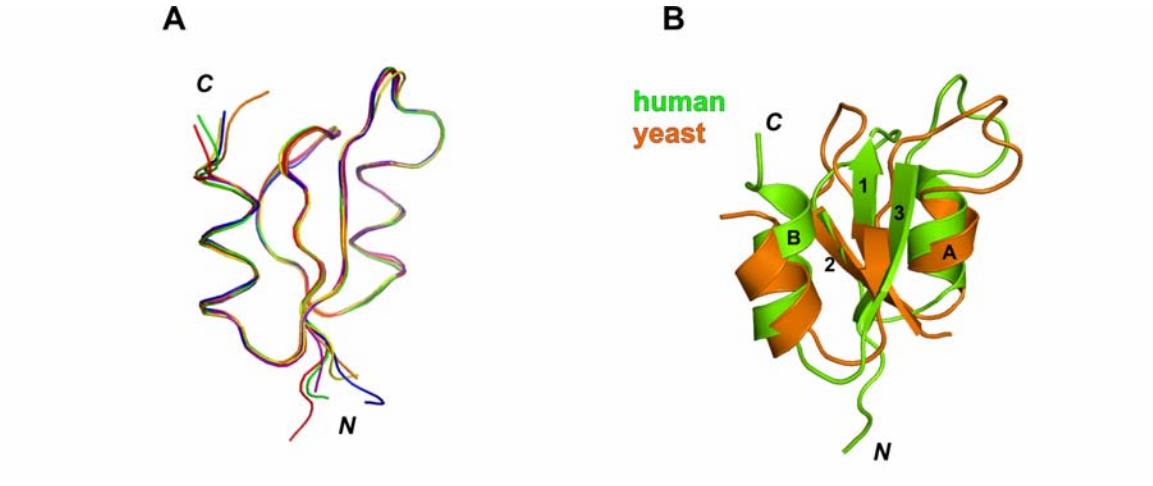
A



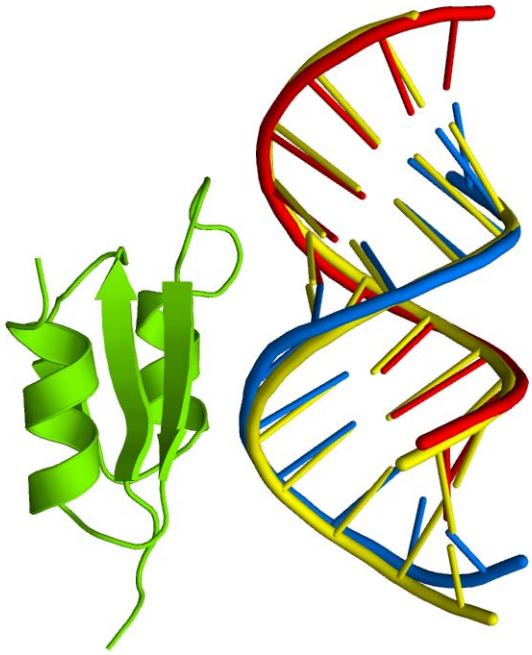
B



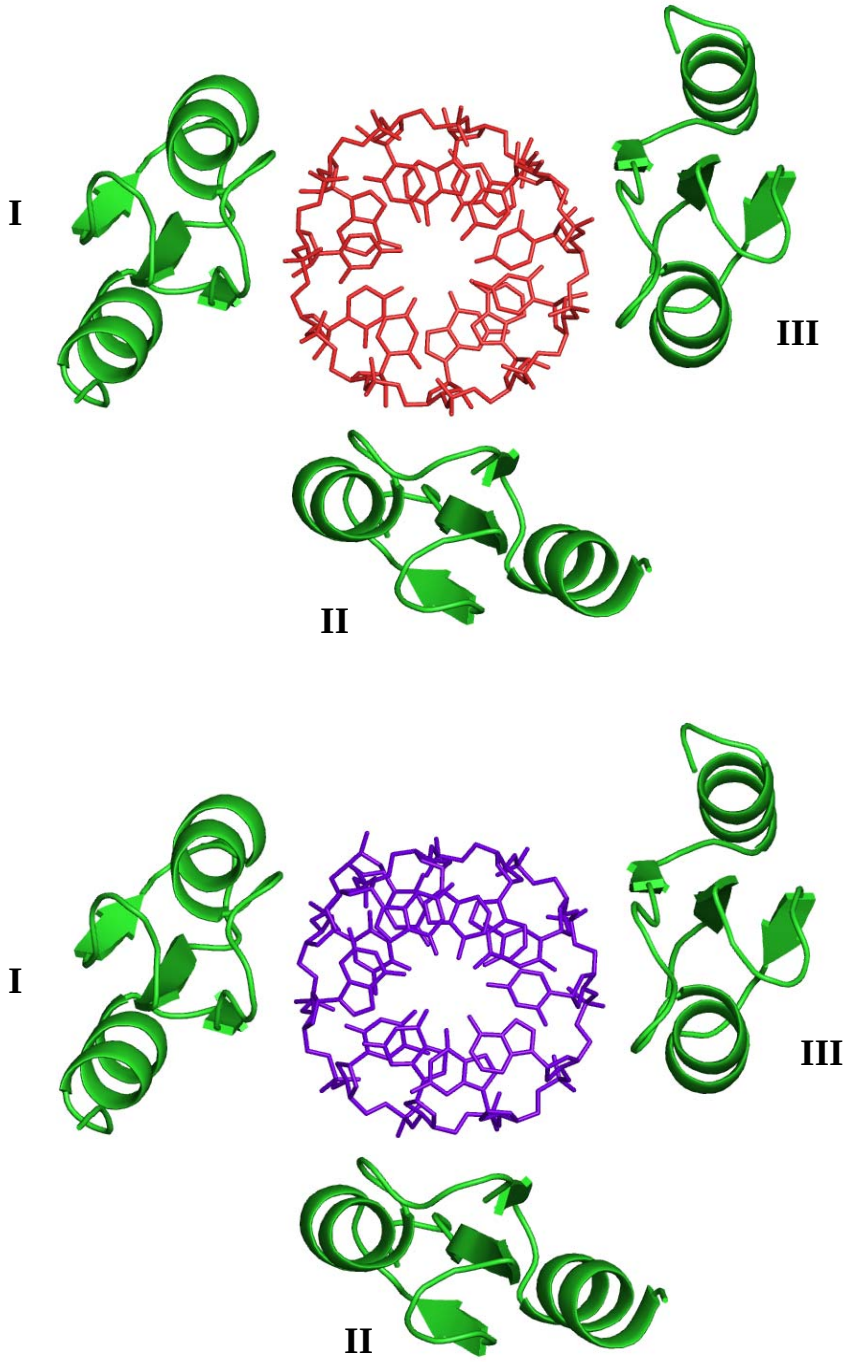
Supplementary Figure 4. (A) A sample diffraction patter image from Hs-HBD/12-mer RNA/DNA crystal. The irregular spot shape is apparent. **(B)** A sample of the omit map after simulated annealing (stereo view). The electron density is contoured at 1.3 σ and overlaid on the region of HBD around the DNA interface.



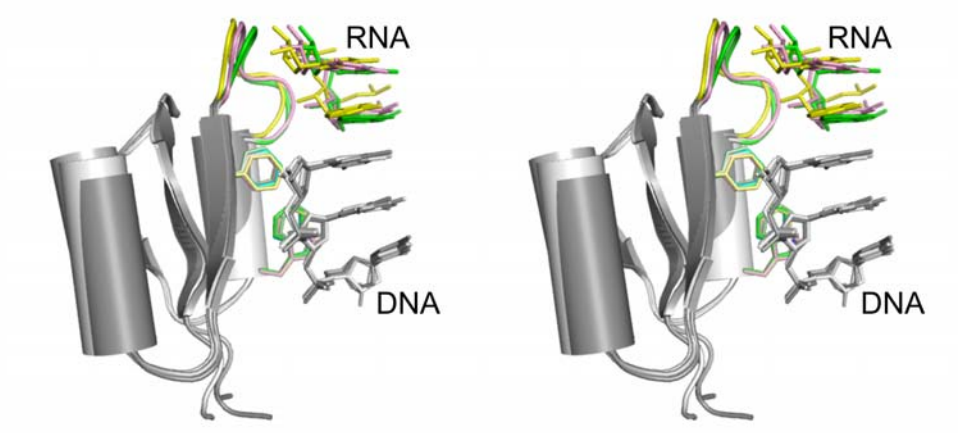
Supplementary Figure 5. Comparison of human and *S. cerevisae* HBD (**A**) Superposition of all six protein molecules present in an asymmetric unit of Hs-HBD - RNA/DNA complex structure. (**B**) Superposition of HBD of *S. cerevisae* RNase H1 determined by NMR (orange, PDB: 1QHK) and human HBD (green, this work).



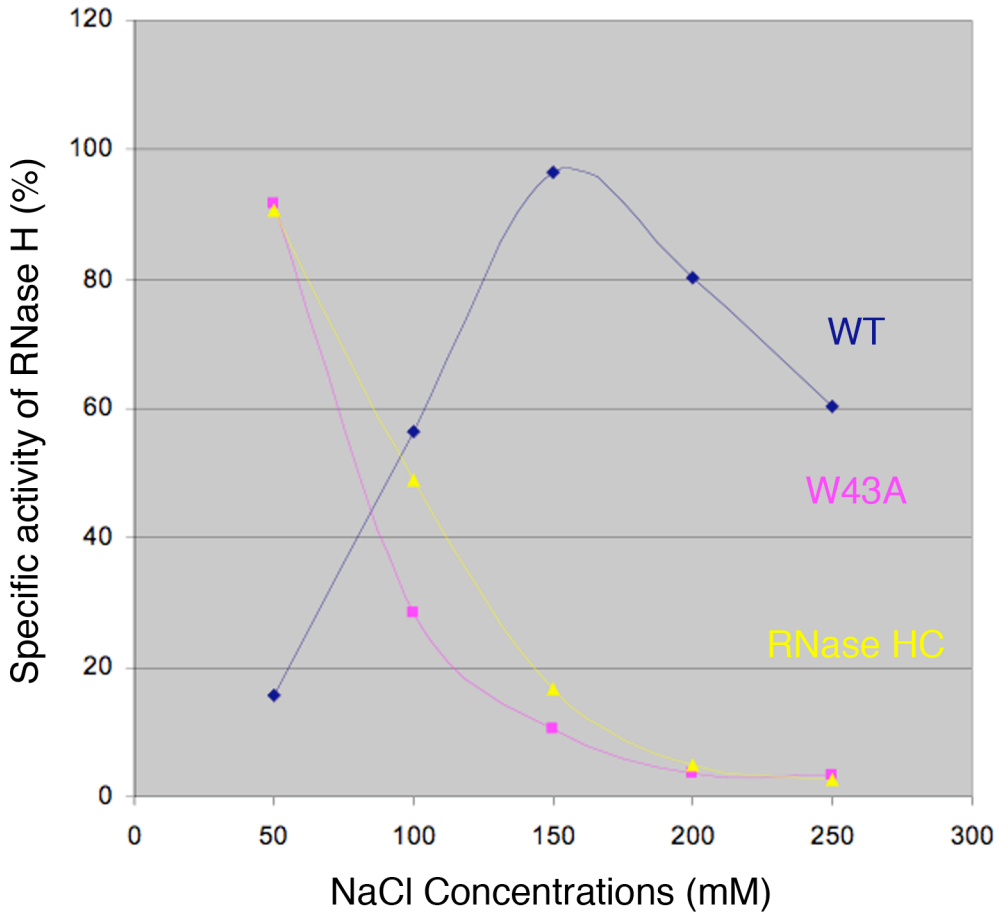
Supplementary Figure 6. Superposition of the HBD-RNA/DNA complex structure (molecule II - green, RNA - red, DNA - blue) with a model of an ideal dsRNA (yellow).



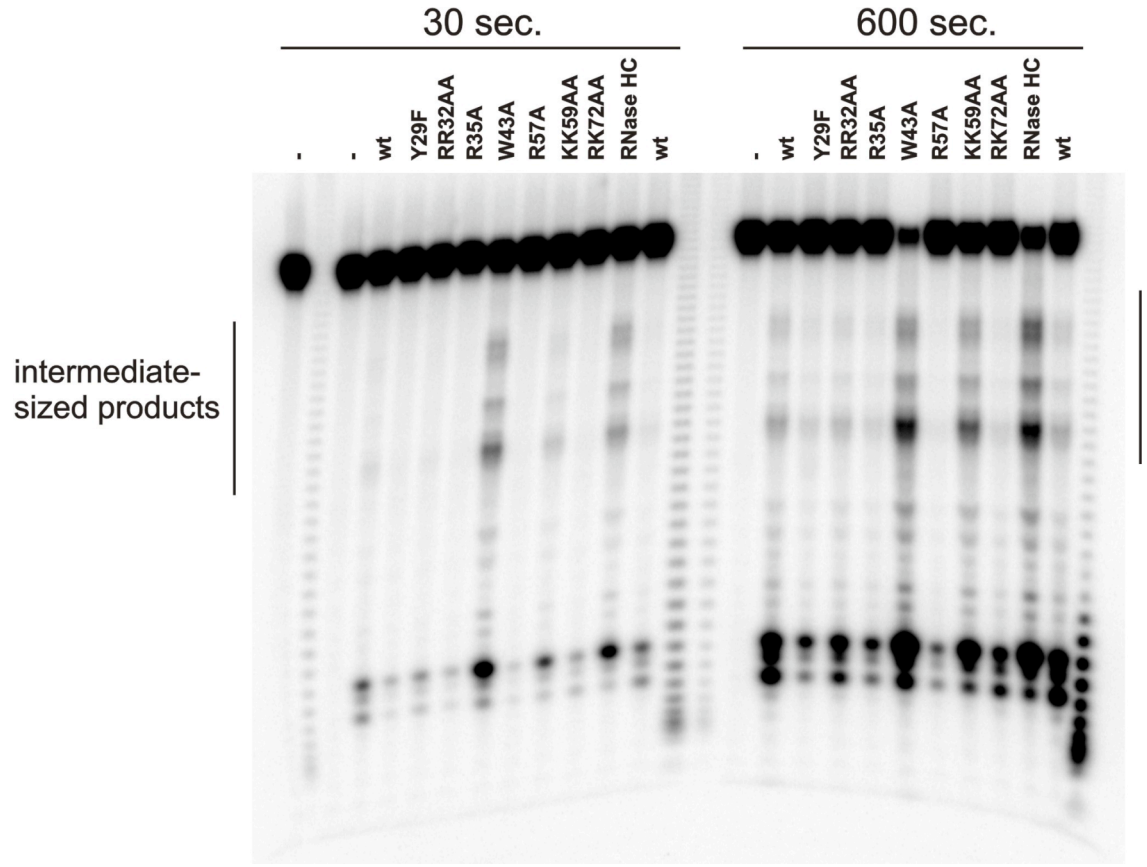
Supplementary Figure 7. Conformation of the nucleic acid in the complex with Hs-HBD, a view down the helical axis. Upper panel shows the RNA strand (in red), the lower panel the DNA (blue). Protein molecules are shown in green.



Supplementary Figure 8. Comparison of three representative HBD molecules and interacting RNA/DNA fragments in an asymmetric unit (stereo view). After superposition of the DNA interface residues (shown in green, yellow and brown sticks), the protein secondary structures and the DNA strand are nearly identical (shown in different shades of grey). The RNA-binding loop and the RNA fragment interacting with it (highlighted in the same color) appear to be co-variants.



Supplementary Figure 9. NaCl concentration dependence of RNase HC (without HBD), W43A (defective HBD) and WT human RNase H1. Each reaction contained 4 pM of enzyme, 0.5 μM ^{32}P-labeled poly-rA)/poly-dT) in 50 mM Tris-HCl pH 7.9, 1 mM DTT, 5 mM $MgCl_2$ and various NaCl in 10 μl. The amount of cleavage products was quantified by liquid scintillation counter.



Supplementary Figure 10. Cleavage of 36bp RNA/DNA by WT and mutant Hs-RNases H1. The experiments were conducted as previously described (Gaidamakov et al., 2005). The amounts of each protein were adjusted to obtain similar degree of digestion (0.1 to 1 nM for full-length RNase H1 and 8 nM of RNase HC). The intermediate-sized products (indicated above) were accumulated in reactions of the W43A and KK59AA mutant RNases H1 as well as by RNase HC.